Comparing methods for modelling spreading cell fronts

Deborah C. Markham[a,*], Matthew J. Simpson[b], Philip K. Maini[a], Eamonn A. Gaffney[a], Ruth E. Baker[a]

[a] *Wolfson Centre for Mathematical Biology, Mathematical Institute, University of Oxford, United Kingdom*
[b] *Mathematical Sciences, Queensland University of Technology, Brisbane, Australia*

Abstract

Spreading cell fronts play an essential role in many physiological processes. Classically, models of this process are based on the Fisher-Kolmogorov equation; however, such continuum representations are not always suitable as they do not explicitly represent behaviour at the level of individual cells. Additionally, many models examine only the large time asymptotic behaviour, where a travelling wave front with a constant speed has been established. Many experiments, such as a scratch assay, never display this asymptotic behaviour, and in these cases the transient behaviour must be taken into account. We examine the transient and asymptotic behaviour of moving cell fronts using techniques that go beyond the continuum approximation via a volume-excluding birth-migration process on a regular one-dimensional lattice. We approximate the averaged discrete results using three methods: (i) mean-field, (ii) pair-wise, and (iii) one-hole approximations. We discuss the performace of these methods, in comparison to the averaged discrete results, for a range of parameter space, examining both the transient and

*Corresponding author: markham@maths.ox.ac.uk

asymptotic behaviours. The one-hole approximation, based on techniques from statistical physics, is not capable of predicting transient behaviour but provides excellent agreement with the asymptotic behaviour of the averaged discrete results, provided that cells are proliferating fast enough relative to their rate of migration. The mean-field and pair-wise approximations give indistinguishable asymptotic results, which agree with the averaged discrete results when cells are migrating much more rapidly than they are proliferating. The pair-wise approximation performs better in the transient region than does the mean-field, despite having the same asymptotic behaviour. Our results show that each approximation only works in specific situations, thus we must be careful to use a suitable approximation for a given system, otherwise inaccurate predictions could be made.
Keywords: travelling front, cell migration, cell proliferation, cancer, wound healing

---

1. Introduction

Advancing fronts of cells are frequently observed experimentally [21, 22, 35, 36]. For example, in Figure 1, we see an advancing front of murine fibroblast 3T3 cells from an *in vitro* experiment [36, 38]. This phenomenon is essential in many physiological processes: embryonic development hinges on the spatial advancement of cells [36], and wounds could not heal without it [21, 22]. Additionally, it is important in tissue engineering [29, 30], which relies on the ability of fronts of cells to move into empty space. Less desirably, moving fronts of cells are a major factor in disease progression, most notably in cancer [1, 11, 37]. An important clinical feature is the sharpness of the

front, which is determined by the relative rates of migration and proliferation; a shallow front can lead to difficulties when surgically removing a tumour [37]. Given their importance biologically, it is hardly surprising that moving cell fronts have been the focus of many mathematical modelling studies.

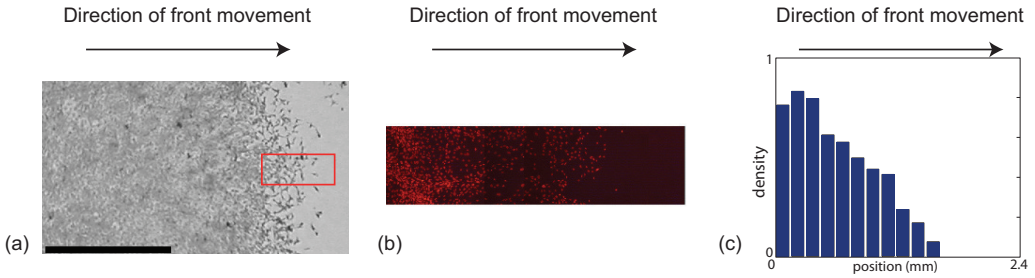


Figure 1: Experimentally observed moving front of murine fibroblast cells. In (a) we see a snapshot of the cells invading the space to the right, whilst in (b), we see the highlighted region from (a) where the cell nuclei have been stained. In (c), we have the calculated density profile showing the shape of the front. The experimental details for producing images such as these can be found in [35].

Classically, advancing fronts of cells have been modelled using the Fisher-Kolmogorov equation [10, 15], which has a travelling wave solution with constant shape and speed. The asymptotic wave speed (as $t \to \infty$), v_f, for initial conditions with compact support, is $2\sqrt{D\lambda}$ where D is the diffusivity of the cells, and λ their effective proliferation rate [25]. Measuring the wavespeed experimentally does not allow us to determine unique values for D and λ, making additional experimental observations necessary [29, 35]. Moreover, even once the travelling wave has been established, the Fisher-Kolmogorov equation, which represents the mean-field behaviour, is not always an accurate representation of the behaviour of a moving front of cells, due to the

stochastic nature of these processes [12, 17]. Thus, whilst it may be possible to fit experimental data to solutions of the Fisher-Kolmogorov model, this does not necessarily lead to accurate parameter estimation; something that is frequently overlooked in models of moving cell fronts [29, 39]. This has led to the development of alternative methods for modelling moving cell fronts, some of which we shall now discuss.

Using agent-based models, each cell is modelled explicitly thus retaining a description of the individual behaviour whilst still enabling observation of the population as a whole [6]. Discrete models have been used to examine moving cell fronts in many areas of cell biology [5, 8, 23]. They are also often used in conjunction with continuum models to provide a multiscale modelling framework [34]. Discrete models are not confined to any particular region of parameter space, but are limited by their computational cost, and lack of analytical tractability. Thus, ideally, we would like to have simpler, more tractable methods approximating the behaviour of moving cell fronts.

When cells proliferate significantly more rapidly than they migrate (λ/D sufficiently large), we expect a sharp front [37] with the region behind the front almost completely filled with cells. Under these conditions, we are able to predict the asymptotic front speed using the one-hole approximation (OHA) [4]. This method uses series expansions to provide a correction term to the front speed for the case without migration, which can be calculated exactly. The OHA agrees well with discrete simulations for sufficiently high λ/D, and can be extended to deal with more than one hole behind the front. However, the method in [4] is only given for constrained systems where a cell either attempts to move or proliferate at every time step, without ever resting.

Additionally, experimental results do not always produce the asymptotic travelling front behaviour. The following three assays highlight some of the different experiments which can be used to obtain data for travelling fronts.

1. A single moving front is allowed to develop over a long period of time (>100 hours) [21, 22]. These experiments are likely to allow for travelling front behaviour to be produced. However, they are not as straightforward to carry out as the same experiment over shorter timescales due to difficulties with keeping the cells alive for long periods of time, and maintaining a constant environment.

2. A single moving front is allowed to develop over a short period of time (<24 hours). The results of these experiments are not on long enough timescales to produce asymptotic travelling front behaviour [30], but are more feasible experimentally.

3. Two opposingly directed fronts come together. For instance, when a thin strip [18, 28, 40, 41] or small hole [41] of cells is removed from a monolayer. In this set-up, the artificially created gap is closed, thus the system may never reach the asymptotic travelling front speed. For example, the protocol in [18] allows between eight and eighteen hours for the scratch to close. Given that typical cell doubling times are of a similar order, we do not expect the asymptotic speed to have been reached before the fronts from either side of the scratch become interwoven. We see an example of this in Figure 2, where a scratch assay is performed with 3T3 cells. Within 30 hours, we see the two fronts meeting.

As many experiments follow the second and third methods, it is often important to be able to predict the transient behaviour as well as the asymptotic speed.

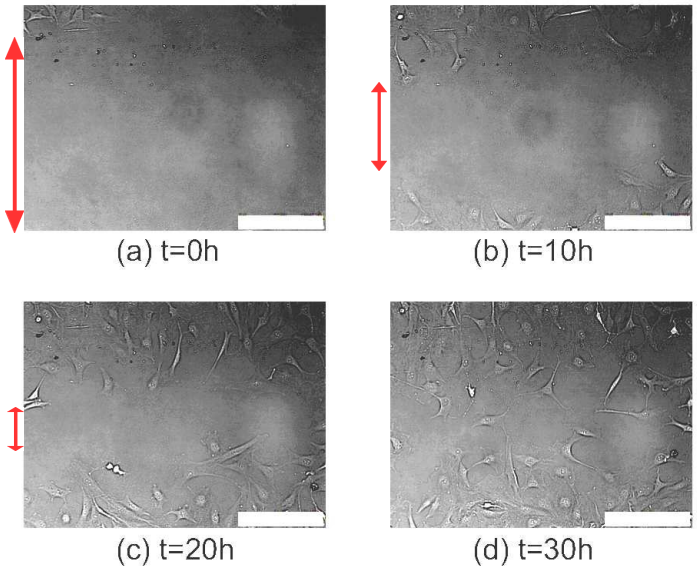


Figure 2: Experimental results of a scratch assay with 3T3 cells. We see that no travelling front is established, as the two fronts quickly collide as they approach from either side of the gap. The white bar corresponds to $250\mu m$. The red arrows give an approximate indication of the width of the scratch at each time point. In Figure (d) there is no arrow as the two fronts have begun to meet in places.

Moment dynamics models incorporate increasingly greater degrees of information into the mean-field model by taking into account the dynamics of cell pairs, triplets, and so forth. Pairwise models are generally the most common, requiring the use of an appropriate closure approximation for any triplet terms in the model [7]. The use of moment dynamics has been well documented in various biological scenarios [2, 3, 16, 26, 31, 32]. Specifi-

cally, in [32], the authors develop a pairwise approximation (PWA), using the Kirkwood Superposition Approximation (KSA) [**13, 14**] as their closure, to describe the behaviour of a moving front. Their work allows the transient behaviour to be examined, shows improvement on the mean-field approximation (MFA), and demonstrates the importance of including spatial correlations in a traditional mean-field model.

In this paper, we examine different methods in detail for a range of proliferation and migration parameters in comparison with averaged discrete results, and discuss which methods are best suited to a given parameter regime and experimental timescale. We begin with a discussion of the discrete Gillespie algorithm used to produce the averaged discrete results. We then discuss the three methods used to approximate the averaged discrete results: the MFA, the PWA, and the OHA. Next, we present results focussing on the transient behaviour, evaluating the relative performance of each method. Following this, we turn to the asymptotic results and examine the methods in a range of parameter space. We conclude by discussing the strengths and weaknesses of the methods in question, and which methods might best suit some specific experimental examples.

2. The methods

In this section we discuss the methods used to model our system. We consider a one-dimensional (1D), volume-excluding process on a regular lattice with a lattice spacing of $\Delta = 1$. Cells move to neighbouring sites at a rate P_m per unit time and proliferate at a rate P_p per unit time.

2.1. The discrete case

For our discrete simulations we have a 1D lattice with $N = 2000$ sites, and we average over 10000 individual realisations to estimate the averaged behaviour. Initially, the l^{th} lattice site is occupied with probability

$$C_l = \begin{cases} 1, & 1 \leq l < x\,, \\ 1 - \left[\dfrac{l-x}{100-x}\right], & x \leq l \leq 100\,, \\ 0, & 100 < l \leq N\,. \end{cases} \tag{2.1}$$

where we can alter the steepness of our initial ramp by varying x. We use a Gillespie algorithm to update our system as in [3], **the algorithm for this being as follows:**

1. **Set $t = 0$.**

2. **Initialize the lattice by generating a uniform random number, r_l, in the interval [0,1], for each lattice site. If $r_l \leq C_l$, an agent is placed at that lattice site.**

3. **With $Q(t)$ being the total number of agents on the lattice, calculate the total propensity function, $a_0 = (P_m + P_p)Q(t)$.**

4. **Calculate the time, τ, to the next event using $\tau = (1/a_0)\mathbf{log}(1/r_1)$, where r_1 is a uniform random number in the interval [0,1].**

5. **Decide which event occurs by calculating $R = a_0 r_2$, where r_2 is another uniform random number in the interval [0,1]. R is used to deduce which event occurs according to the following:**

- **If $R \in [0, P_m Q(t))$, a movement event will occur. An agent is chosen at random, and one of its neighbours is also chosen at random as the target site for a movement event. If the target site is empty, the chosen agent moves to that site, otherwise the event is aborted.**
- **If $R \in [P_m Q(t), (P_m + P_p)Q(t)]$, a proliferation event will occur. An agent is chosen at random, and one of its neighbours is also chosen at random as the target site for a proliferation event. If the target site is empty, a daughter cell is placed in it, otherwise the event is aborted.**

6. **Update $Q(t)$ depending on which event, if any, occurred.**
7. **Update time by setting $t \to t + \tau$.**
8. **Repeat from step 3 until the stipulated final time is reached.**

We apply reflecting boundary conditions at $l = 1$ and $l = N$. To determine the front speed, v, we track the location where averaged cell density is 0.5, and calculate the average velocity from this information. Once v is no longer changing with time, we have reached the asymptotic travelling front speed, v_f.

2.2. The mean-field approximation

To derive the MFA we consider the occupancy of each site of the lattice. The average occupancy of the l^{th} lattice site is given by $C_l \in [0, 1]$. We use k-point distribution functions, $\rho^{(k)}$, to **derive the MFA and PWA. The k-point distribution functions describe the probability that k-tuplets**

of sites have given occupancies. The one-point distribution function, $\rho^{(1)}$, gives the averaged occupancy of the site in question. Thus we have

$$\rho^{(1)}(A_l) = C_l, \tag{2.2}$$

$$\rho^{(1)}(0_l) = 1 - C_l, \tag{2.3}$$

where A_l and 0_l indicate that site l is occupied by A or unoccupied, respectively. To provide a measure of the occupancy dependence of two given sites we use correlation functions, as defined in [3, 19, 20]:

$$F_{\sigma_l,\sigma_m}(l,m) \coloneqq \frac{\rho^{(2)}(\sigma_l,\sigma_m)}{\rho^{(1)}(\sigma_l)\rho^{(1)}(\sigma_m)}, \tag{2.4}$$

where the state of the site is given by σ_l, which is either 0 or A. If lattice site occupancies are independent, we have $F_{\sigma_l,\sigma_m}(l,m) \equiv 1$. **Correlation functions can be related by using conservation expressions. For example, we can write the conservation equation**

$$\rho^{(2)}(A_l, A_m) + \rho^{(2)}(A_l, 0_m) = \rho^{(1)}(A_l). \tag{2.5}$$

Using the definition of a correlation function, we have

$$C_l C_m F_{A,A}(l,m) + C_l(1 - C_m)F_{A,0}(l,m) = C_l, \tag{2.6}$$

and thus we can express one correlation function in terms of another:

$$F_{A,0}(l,m) = \frac{1 - C_m F_{A,A}(l,m)}{1 - C_m}. \tag{2.7}$$

We now turn to the evolution of the 1-point distribution function:

$$\frac{d\rho^{(1)}(A_l)}{dt} = \underbrace{\frac{P_m}{2}\left[\rho^{(2)}(A_{l-1},0_l)+\rho^{(2)}(0_l,A_{l+1})-\rho^{(2)}(A_l,0_{l+1})-\rho^{(2)}(0_{l-1},A_l)\right]}_{\text{movement in and out of site } l} + \underbrace{\frac{P_p}{2}\left[\rho^{(2)}(A_{l-1},0_l)+\rho^{(2)}(0_l,A_{l+1})\right]}_{\text{proliferation into site } l}. \tag{2.8}$$

We rewrite equation (2.8) using the correlation functions and conservation equations to obtain

$$\begin{aligned}\frac{dC_l}{dt} =& \frac{P_m}{2}\left[C_{l-1}-2C_l+C_{l+1}\right] \\ &+\frac{P_p}{2}\left\{C_{l-1}\left[1-C_lF(l,l-1)\right]+C_{l+1}\left[1-C_lF(l,l+1)\right]\right\},\end{aligned} \tag{2.9}$$

where we have set $F_{A,A} = F$ for notational simplicity. **We note that the movement term does not depend on correlations, as all correlation containing terms cancel [32].** For the standard MFA, we assume lattice site occupancies are independent, thus $F(l,m) \equiv 1$ for all l, m. Therefore, we have the following equation for the evolution of the density of a given lattice site:

$$\begin{aligned}\frac{dC_l}{dt} =& \frac{P_m}{2}\left[C_{l-1}-2C_l+C_{l+1}\right] \\ &+\frac{P_p}{2}\left[C_{l-1}+C_{l+1}\right]\left[1-C_l\right].\end{aligned} \tag{2.10}$$

To make predictions using the MFA, we solve the system given by equation (2.10) numerically using a fourth order Runge-Kutta method (RK4) [27] with a constant time-step of $\delta t = 0.1$. Smaller timesteps were tested to confirm this was an appropriate choice. We calculate v and v_f in the same manner as in our discrete model, whereby we track the location of $C_l = 0.5$.

We have reflecting boundary conditions and initial conditions of the form of equation (2.1).

2.3. Pair-wise approximation

For the PWA, we no longer make the usual assumption that the occupancies of pairs of sites are independent, thus we do not set $F(l,m) \equiv 1$ in equation (2.9). To determine the evolution of our correlation functions, we turn to the 2-point distribution functions. First we consider where the two sites in question are nearest neighbours:

$$\begin{aligned}\frac{d\rho^{(2)}(A_l, A_{l+1})}{dt} = &\underbrace{\frac{P_m}{2}\left[\rho^{(3)}\left(A_{l-1}, 0_l, A_{l+1}\right) + \rho^{(3)}\left(A_l, 0_{l+1}, A_{l+2}\right)\right]}_{\text{movement into sites } l \text{ and } l+1} \\ &- \underbrace{\frac{P_m}{2}\left[\rho^{(3)}\left(0_{l-1}, A_l, A_{l+1}\right) + \rho^{(3)}\left(A_l, A_{l+1}, 0_{l+2}\right)\right]}_{\text{movement out of sites } l \text{ and } l+1} \\ &+ \underbrace{\frac{P_p}{2}\left[\rho^{(3)}\left(A_{l-1}, 0_l, A_{l+1}\right) + \rho^{(3)}\left(A_l, 0_{l+1}, A_{l+2}\right)\right]}_{\text{proliferation from neighbours}} \\ &+ \underbrace{\frac{P_p}{2}\left[\rho^{(2)}\left(0_l, A_{l+1}\right) + \rho^{(2)}(A_l, 0_{l+1})\right]}_{\text{proliferation from each other}}.\end{aligned} \tag{2.11}$$

Next, we consider the evolution of the 2-point distribution function where

the two sites in question are not nearest neighbours, thus $|l-m|>1$:

$$\begin{aligned}\frac{d\rho^{(2)}(A_l,A_m)}{dt} =& \underbrace{\frac{P_m}{2}\left[\rho^{(3)}(A_{l-1},0_l,A_m)+\rho^{(3)}(0_l,A_{l+1},A_m)\right]}_{\text{movement into site } l}\\ &-\underbrace{\frac{P_m}{2}\left[\rho^{(3)}(0_{l-1},A_l,A_m)+\rho^{(3)}(A_l,0_{l+1},A_m)\right]}_{\text{movement out of site } l}\\ &+\underbrace{\frac{P_m}{2}\left[\rho^{(3)}(A_l,A_{m-1},0_m)+\rho^{(3)}(A_l,0_m,A_{m+1})\right]}_{\text{movement into site } m}\\ &-\underbrace{\frac{P_m}{2}\left[\rho^{(3)}(A_l,0_{m-1},A_m)+\rho^{(3)}(A_l,A_m,0_{m+1})\right]}_{\text{movement out of site } m}\\ &+\underbrace{\frac{P_p}{2}\left[\rho^{(3)}(A_{l-1},0_l,A_m)+\rho^{(3)}(0_l,A_{l+1},A_m)\right]}_{\text{proliferation into site } l}\\ &+\underbrace{\frac{P_p}{2}\left[\rho^{(3)}(A_l,A_{m-1},0_m)+\rho^{(3)}(A_l,0_m,A_{m+1})\right]}_{\text{proliferation into site } m}.\end{aligned} \tag{2.12}$$

In equations (2.11) and (2.12), we use conservation expressions for the 3-point distribution functions, as in [3, 24, 32], to eliminate some of the 3-point distribution function terms. Where this is not possible, we close using the KSA**, which is given by the following equation:**

$$\rho^{(3)}(\sigma_l,\sigma_m,\sigma_n)=\frac{\rho^{(2)}(\sigma_l,\sigma_m)\rho^{(2)}(\sigma_l,\sigma_n)\rho^{(2)}(\sigma_m,\sigma_n)}{\rho^{(1)}(\sigma_l)\rho^{(1)}(\sigma_m)\rho^{(1)}(\sigma_n)}. \tag{2.13}$$

We relate the 2-point distribution functions to the correlation functions using equation (2.4) to obtain equations for the evolution of the correlation

functions. The evolution of $F(l, l+1)$ is given by

$$
\begin{aligned}
F'(l,l+1) &= -F(l,l+1)\left[\frac{C'_{l+1}}{C_{l+1}} + \frac{C'_l}{C_l}\right] \\
&+ \frac{P_m}{2}\left[\frac{C_{l-1}}{C_l}F(l-1,l+1) + \frac{C_{l+2}}{C_{l+1}}F(l,l+2) - 2F(l,l+1)\right] \\
&+ \frac{P_p}{2}\left[\frac{1}{C_l} + \frac{1}{C_{l+1}} - 2F(l,l+1)\right. \\
&+ \frac{C_{l-1}}{C_l(1-C_l)}\left[1 - C_lF(l,l+1)\right]\left[1 - C_lF(l,l-1)\right]F(l-1,l+1) \\
&+ \left.\frac{C_{l+2}}{C_{l+1}(1-C_{l+1})}\left[1 - C_{l+1}F(l,l+1)\right]\left[1 - C_{l+1}F(l+1,l+2)\right]F(l,l+2)\right],
\end{aligned}
\tag{2.14}
$$

where prime denotes differentiation with respect to time. For any general distance $F(l,m)$, where $|l-m| > 1$, we obtain

$$
\begin{aligned}
F'(l,m) &= -F(l,m)\left[\frac{C'_l}{C_l} + \frac{C'_m}{C_m}\right] + \frac{P_m}{2}\left[\frac{C_{l-1}}{C_l}F(l-1,m) + \frac{C_{l+1}}{C_l}F(l+1,m)\right. \\
&+ \left.\frac{C_{m-1}}{C_m}F(l,m-1) + \frac{C_{m+1}}{C_m}F(l,m+1) - 4F(l,m)\right] \\
&+ \frac{P_p}{2}\left[\frac{C_{l-1}}{C_l(1-C_l)}\left[1 - C_lF(l,l-1)\right]\left[1 - C_lF(l,m)\right]F(m,l-1)\right. \\
&+ \frac{C_{l+1}}{C_l(1-C_l)}\left[1 - C_lF(l,l+1)\right]\left[1 - C_lF(l,m)\right]F(m,l+1) \\
&+ \frac{C_{m+1}}{C_m(1-C_m)}\left[1 - C_mF(l,m)\right]\left[1 - C_mF(m,m+1)\right]F(l,m+1) \\
&+ \left.\frac{C_{m-1}}{C_m(1-C_m)}\left[1 - C_mF(l,m)\right]\left[1 - C_mF(m,m-1)\right]F(l,m-1)\right].
\end{aligned}
\tag{2.15}
$$

We solve these equations numerically using an RK4 algorithm with a constant time step, δt. For each set of parameters, we test smaller timesteps than the one used to confirm the results are visually indistinguishable. We calculate v and v_f in the same way as for the MFA, and we again have

reflecting boundary conditions and initial conditions of the form of equation (2.1). As discussed in [32], initial conditions such as these lead to difficulties as the correlation functions are unbounded when $C_l = 1$ or $C_l = 0$. This issue can be resolved by using a hybrid approach whereby we use the PWA in regions where $\epsilon < C < (1 - \epsilon)$ and the MFA elsewhere. In this case, we set $\epsilon = 1 \times 10^{-10}$. We also need to choose a truncation value for $F(l, m)$ with $m = l + 1, l + 2, l + 3, ..., l + M$. We truncate at $M = 10$, and test higher values of M to confirm our truncation choice (results not shown).

2.4. One hole approximation

The OHA estimates the asymptotic front speed in situations where we assume there is only one hole (unoccupied lattice site) behind the front. In this section we show how to calculate v_f using the OHA, relaxing the assumption made in [4] that $P_p + P_m = 1$. We know that if $P_m = 0$, $v_f = P_p/2$ and there will be no holes behind the moving front. For P_m small, there will be a correction factor to v_f. In this approximation, we assume there will never be more than one hole behind the leading edge of the front. We define the one-hole states, whereby the hole is in the n^{th} position behind the leading cell, in the following way:

- $|0\rangle$ = (...11111000...);
- $|1\rangle$ = (...11101000...);
- $|2\rangle$ = (...11011000...);
- $|3\rangle$ = (...10111000...), etc.

Note that we always define states in a frame moving with respect to the front. We only allow transitions between these states. The transitions and their associated probabilities $W_{ij} \equiv W(|i\rangle \to |j\rangle)$ are:

- $W_{00} = \underbrace{P_p/2}_{\text{proliferation forwards}},$
- $W_{01} = \underbrace{P_m/2}_{\text{migration forwards}},$
- $W_{10} = \underbrace{P_p}_{\text{proliferation of cells either side of hole}} + \underbrace{P_m/2}_{\text{migration of end cell backwards}},$
- $W_{12} = \underbrace{P_p/2}_{\text{proliferation of end cell forwards}} + \underbrace{P_m/2}_{\text{migration of second to end cell forwards}},$
- $W_{n0} = \underbrace{P_p}_{\text{proliferation of cells either side of hole}},$
- $W_{n,n-1} = \underbrace{P_m/2}_{\text{migration from ahead of hole}}, \qquad n > 1,$
- $W_{n,n+1} = \underbrace{P_p/2}_{\text{proliferation of end cell}} + \underbrace{P_m/2}_{\text{migration from behind hole}}, \qquad n > 1.$

Next, we consider the probability of being in a given state, p_n. **As we are only considering the asymptotic behaviour, we know that transitions in and out of a state must be equal, leading to the following equations**:

$$p_0 W_{01} = \sum_{n=1}^{\infty} p_n W_{n0}, \tag{2.16}$$

$$p_1(W_{10} + W_{12}) = p_0 W_{01} + p_2 W_{21}, \tag{2.17}$$

$$p_n(W_{n0} + W_{n,n-1} + W_{n,n+1}) = p_{n-1} W_{n-1,n} + p_{n+1} W_{n+1,n}, \quad n > 1. \tag{2.18}$$

Using the transition probabilities, we can rewrite these as

$$\frac{P_m}{2}p_0 = \left(P_p + \frac{P_m}{2}\right)p_1 + P_p\sum_{n=2}^{\infty}p_n, \tag{2.19}$$

$$\left(\frac{3P_p}{2} + P_m\right)p_1 = \frac{P_m}{2}p_0 + \frac{P_m}{2}p_2, \tag{2.20}$$

$$\left(\frac{3P_p}{2} + P_m\right)p_n = \frac{P_p + P_m}{2}p_{n-1} + \frac{P_m}{2}p_{n+1}, \quad n > 1. \tag{2.21}$$

For $n > 1$, we follow the approach of [4] and assume the ansatz $p_n = a^{n-1}p_1$. We insert this ansatz into equation (2.21) to obtain

$$\frac{P_m}{2}a^2 - \left(\frac{3P_p}{2} + P_m\right)a + \frac{P_p + P_m}{2} = 0. \tag{2.22}$$

Solving this quadratic in a gives

$$a = \frac{3P_p/2 + P_m \pm \sqrt{9P_p^2/4 + 2P_pP_m}}{P_m}. \tag{2.23}$$

We know that as n increases p_n decreases, therefore we take only the negative square root. Assuming that P_m is small relative to P_p, we can use a series expansion to obtain

$$a = \frac{1}{3} + \frac{4}{27}\left(\frac{P_m}{P_p}\right) - \frac{16}{243}\left(\frac{P_m}{P_p}\right)^2 + \frac{80}{2187}\left(\frac{P_m}{P_p}\right)^3 + \mathcal{O}\left(\frac{P_m}{P_p}\right)^4. \tag{2.24}$$

Note that even if we enforce $P_p + P_m = 1$, this is different from the expression in [4]. Having relaxed this assumption earlier, the expansion was performed on a different expression, thus leading to a different overall approximation.

We substitute $p_2 = ap_1$ into equation (2.20) and use $\sum_{n=2}^{\infty}p_n = 1 - p_0 - p_1$ in equation (2.19) resulting in a pair of equations for p_0 and p_1. We solve these simultaneous equations to obtain

$$p_0 = \frac{3P_pP_m + 4P_p^2 - P_m\sqrt{P_p(9P_p + 8P_m)}}{2(P_pP_m - P_m^2 + 2P_p^2)}, \tag{2.25}$$

$$p_1 = \frac{6P_p^2 + 7P_pP_m - P_m\sqrt{P_p(9P_p + 8P_m)} + 2P_p\sqrt{P_p(9P_p + 8P_m)}}{2(P_pP_m - P_m^2 + 2P_p^2)}. \quad (2.26)$$

Knowing p_0, it is possible for us to approximate v_f. We know that if holes are opening up behind the front, v_f will be smaller. When $P_m = 0$, $v_f = P_p/2$. For non-zero P_m, there will sometimes be movement forward at a rate $P_m/2$, thus, if there are no holes, $v_f = (P_p + P_m)/2$. There will also be movement back into holes thus we reduce the front speed by taking into account those agents moving back into holes:

$$v_f = \frac{P_p + P_m}{2} - \frac{P_m p_1(P_p, P_m)}{2}. \quad (2.27)$$

Performing a series expansion on p_1, we thus obtain the following approximate expression for the front speed:

$$v_f = \frac{P_p + P_m}{2} - \frac{1}{6}\frac{P_m^2}{P_p} + \frac{5}{54}\frac{P_m^3}{P_p^2} - \frac{340 - 162P_p}{3888}\frac{P_m^4}{P_p^3} + \mathcal{O}(P_m^5). \quad (2.28)$$

Therefore, we estimate v_f to be 0.498 when $P_p = 0.9$ and $P_m = 0.1$, agreeing with work in [4].

3. Results

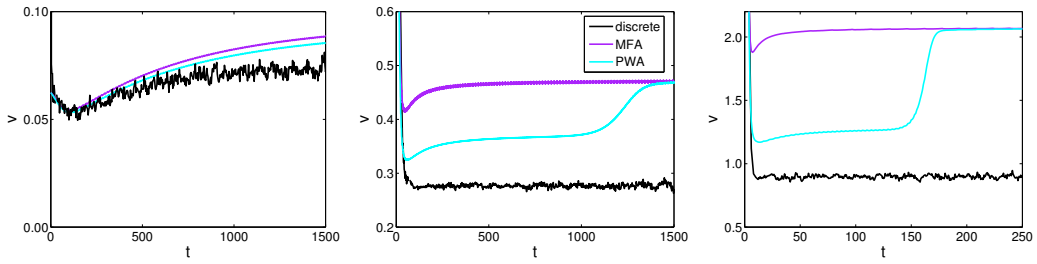
3.1. Transient behaviour

In practice, we must wait a sufficient duration of time to observe the asymptotic speed, which in theory is only fully established as $t \to \infty$. In Figure 3, we see this for the three cases in which we can examine the transient behaviour: the discrete, MFA and PWA. The OHA is only capable of predicting the asymptotic front speed, which is not ideal as this may not always be reached in reality, depending on the experimental conditions

[18, 28, 40, 41]. We see, in Figure 3, that the time taken to reach the asymptotic travelling front speed varies depending on the model chosen, and the parameter values. For sufficiently low P_p/P_m, the MFA and PWA are both in suitable agreement with the averaged discrete results. For larger P_p/P_m, the discrete model reaches the asymptotic speed noticeably faster than the MFA. The PWA takes significantly longer and changes very slowly for a long period before more rapidly adjusting to the asymptotic speed. The asymptotic speed, v_f, is identical for the MFA and PWA in all cases, with this value being higher than the averaged discrete result for large P_p/P_m. The asymptotic speed is identical for the MFA and PWA due to the fact we have used a hybrid PWA whereby we use the MFA in regions where $C < \epsilon$ and $C > 1 - \epsilon$. Changing ϵ will slightly shift the point at which the asymptotic speed is reached; a larger value of ϵ leads to the asymptotic speed being reached more quickly (results not shown).

Additionally, we examine the effects of altering the steepness of the initial conditions by varying x in Equation (2.1). Altering the initial conditions, as long as compact support is maintained, does not affect the asymptotic results. However, different initial conditions have an impact on the transient behaviour. We see, in Figure 4, that a steeper ramp (which corresponds to a higher value of x in Equation (2.1)) leads to a lower initial speed. The MFA and PWA also lie slightly closer to the averaged discrete behaviour at early times when the initial conditions are steeper.

We compare the transient density profiles at various times in Figure 5. From this we see that the PWA better predicts the averaged discrete be-



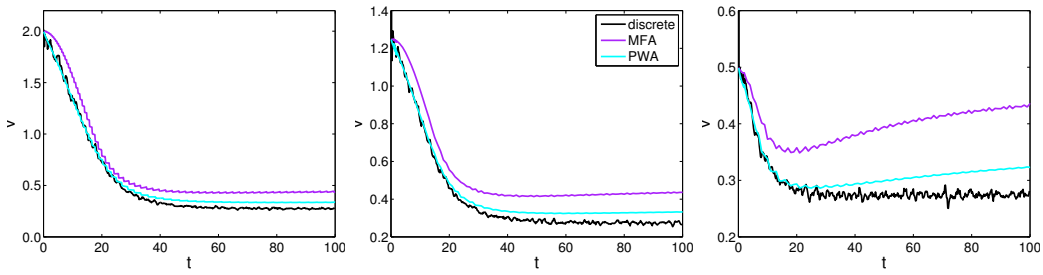
(a) $P_p/P_m = 5 \times 10^{-3}$ (b) $P_p/P_m = 10^{-1}$ (c) $P_p/P_m = 1$

Figure 3: The transient front speed, v, varies over time, eventually reaching its asymptotic travelling speed, v_f. We see this behaviour in the three different cases for which we can examine the transient behaviour: the discrete, MFA and PWA. The behaviour depends on the parameters chosen, and we compare three P_p/P_m ratios (for all cases $P_m = 1$ and P_p is varying). For low P_p/P_m, we notice that all three methods are in relatively good agreement. For larger P_p/P_m the MFA and PWA begin to deviate from the averaged discrete results, generally tending to a higher asymptotic speed. In all cases, the MFA and PWA eventually reach the same travelling speed, although the PWA takes significantly longer. The time taken for the PWA to reach the same speed as the MFA decreases as P_p/P_m increases, as we see by comparing (b) and (c).

haviour than does the MFA. Thus, whilst the asymptotic behaviour might be the same, the PWA is more accurate at predicting the average transient behaviour of the system when P_p is sufficiently low relative to P_m. **We also note, by comparing the top and bottom rows of Figure 5, that a steeper initial slope leads to the MFA and PWA better approximating the averaged discrete behaviour in the transient region.**

3.2. Asymptotic behaviour

As $t \to \infty$, all cases predict a travelling front with a constant speed. **We relax the assumption in [4], no longer requiring that** $P_p + P_m = 1$**.**



(a) $x = 20$ (b) $x = 50$ (c) $x = 80$

Figure 4: **Adjusting the steepness of the initial conditions affects the transient behaviour. We see that a steeper slope (given by a higher value of x) leads to the initial speed being lower, and the PWA and MFA providing a slightly closer agreement with the averaged discrete behaviour. In all cases, we see that the PWA provides a significant improvement to the MFA at predicting the averaged discrete behaviour in the transient region.**

This allows us to look at a more relative measure, the ratio of proliferation to migration rates, P_p/P_m. We keep $P_m = 1$ and allow P_p to vary, examining the resulting behaviour in Figure 6(a). We see that the OHA provides a good estimate of v_f for high enough P_p/P_m, but diverges when P_p/P_m is reduced beyond a certain level. We also look at the average number of holes behind the front in the discrete case (Figure 6(b)) for the same parameter range, noting that this corresponds well with the predictive power of the OHA: as we move beyond one hole on average, the OHA breaks down. The break-down of the OHA can also be attributed to the assumption that P_m is small relative to P_p which ceases to hold as we increase P_m and decrease P_p. The MFA prediction for v_f (which is identical to the PWA predictions) agrees well with the averaged discrete case for low P_p but is not a good predictor when P_p is large. For a region in the middle, neither

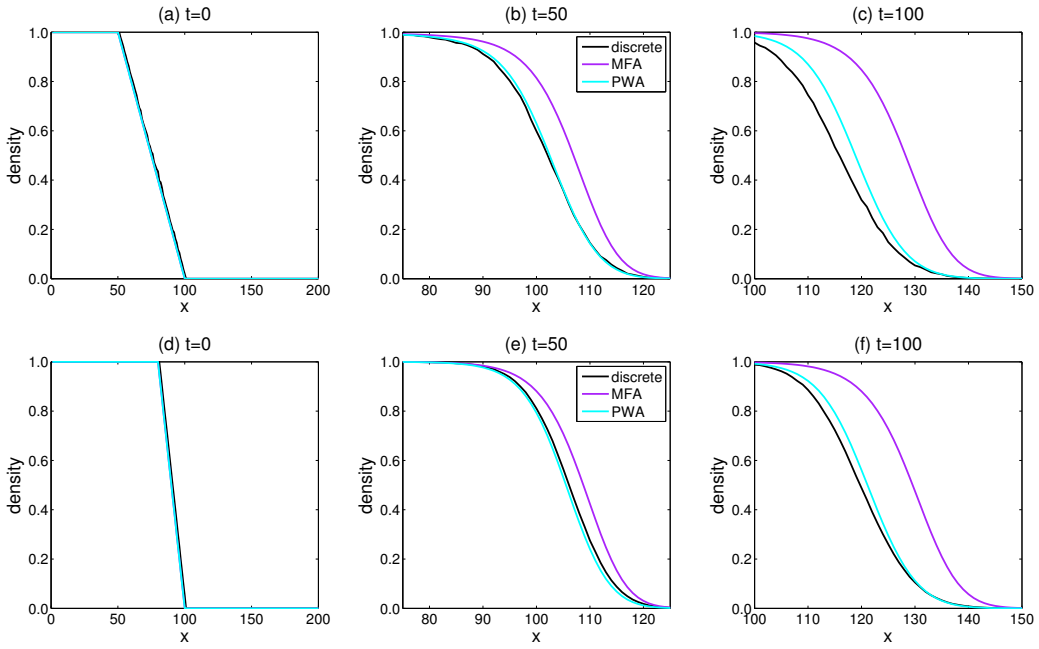
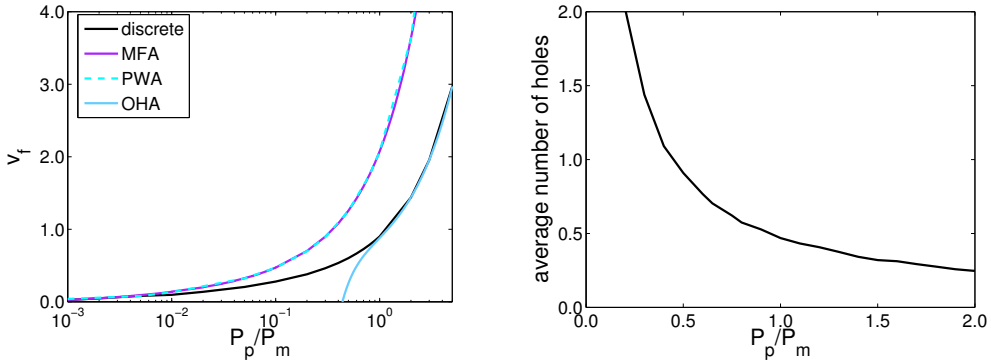


Figure 5: Density plots for early times show that the PWA provides a closer approximation of the averaged discrete behaviour than does the MFA. The parameters in this case are $P_m = 1, P_p = 0.1$. **The top row are for initial conditions whereby $x = 50$, and the bottom row results are for $x = 80$. We see that a steeper ramp leads to improved performance of the MFA and PWA at approximating the transient averaged discrete behaviour.**



(a) Asymptotic speed

(b) Holes behind the front

Figure 6: The asymptotic travelling wave speed varies with the parameters. In (a), we see that the OHA performs well at approximating the averaged discrete behaviour when P_p/P_m is sufficiently large, and we look at the average number of holes for the discrete case in (b). The MFA and PWA, which are identical, perform well when P_p/P_m is sufficiently small.

the MFA nor the OHA provides a suitable approximation to the averaged discrete behaviour.

4. Discussion

Mathematical models are often used in conjunction with experimental data to examine moving cell fronts, in the hope of determining information such as the mechanisms driving the movement of the front. We have compared different methods for modelling the transient and asymptotic moving cell front behaviour in a range of parameter space. These results are summarized in Table 1. Many models in the past have focussed on predicting the asymptotic behaviour, whereby a front travelling at constant speed has been established. Whilst this is appropriate in some cases, there are many exper-

	Small P_p/P_m		Intermediate P_p/P_m		Large P_p/P_m	
	Short	Long	Short	Long	Short	Long
MFA	✓	✓	✗	✗	✗	✗
PWA	✓	✓	✗	✗	✗	✗
OHA	✗	✗	✗	✗	✗	✓
Discrete	✓	✓	✓	✓	✓	✓

Table 1: A description of which methods work under specific circumstances, **where long and short refer to the asymptotic and transient behaviour, respectively**. The best method for given conditions is highlighted in green.

imental situations in which the transient behaviour is key. Not only does the final behaviour take some time in practice to achieve, but some experiments are performed in such a manner that we may never observe the final behaviour. For instance, in a scratch or punch-hole assay, the cells will be encroaching from multiple directions, resulting in the fronts interacting and the unoccupied region closing before asymptotic behaviour can be reached [18, 28, 40]. A major drawback of the OHA is its inability to predict the transient behaviour; it is only able to predict the asymptotic speed. The MFA and PWA can both be used for transient data, with the PWA giving improved results thus making it preferable for transient behaviour.

For asymptotic behaviour, the method best suited to a given situation depends on the relative rates of movement and proliferation. In regions where P_p/P_m is small ($P_p/P_m < 0.01$), we have a shallow front [37] with many holes behind the leading cell, thus the OHA diverges substantially from the averaged discrete results and cannot be successfully used. The MFA and PWA are better suited to modelling cells in this region of parameter space.

As these two methods give the same result asymptotically, it is most sensible to use the MFA for asymptotic predictions as this is a far simpler model to implement. In the experimental results in Figure 1, we notice a large number of holes behind the front, thus we expect this cell line [36, 38] to be best approximated by the MFA asymptotically, and by the PWA in the transient region.

For cell lines where P_p/P_m is relatively large ($P_p/P_m > 0.8$), there will be fewer holes behind the sharp front [37], thus the OHA is the method best suited to asymptotic predictions in this region, as indicated by its excellent agreement with the averaged discrete results. The MFA and PWA do not perform as well in this region. We expect cell lines with low rates of migration relative to proliferation to lie in this region of parameter space. For example, in [33], a pair-wise model developed for uniform initial conditions [3], was used to determine P_p and P_m for a breast cancer cell line, MDA MB 231. With $P_p = 0.069$ and $P_m = 0.04$, we have a ratio, $P_p/P_m = 1.7$, which lies within the range for which the OHA is best suited for asymptotic predictions. Thus for MDA MB 231 and similar cell lines, we should use the OHA if we only need information about the asymptotic behaviour.

For intermediate P_p/P_m ($0.01 < P_p/P_m < 0.8$), none of the methods provide a good approximation to the averaged discrete results in the long term, and the PWA becomes less successful in the transient region. Thus developing a model that accurately approximates the averaged discrete behaviour in this region requires further investigation. By extending the OHA to higher numbers of holes, as in [4], we can improve upon this approximation, but at an increasing level of complexity for every additional hole added. Similarly,

the PWA can be extended to triplets and so forth to better predict the transient behaviour in a wider range of parameter space. Again, this becomes increasingly complex as more information is incorporated.

In future, one could consider the inclusion of death in models of advancing cell fronts. It is fairly straightforward to include death in the MFA and PWA, and this has been done in previous work [32]. The OHA can also be extended to include death. However, if death rates are too high, relative to proliferation, we expect the OHA to break down due to there being a large number of holes behind the front. Additionally, we have only considered a 1D problem but the method can be extended to higher dimensions in future in order to be more biologically relevant. Some work has been done on predicting the asymptotic front speed by following the leading cell in a similar way to the OHA for ecological systems [9].

We have demonstrated the relative merits of the MFA, PWA and OHA in comparison with averaged discrete results for transient and asymptotic behaviour in a broad range of parameter space. Our results demonstrate that it is essential to choose the most appropriate modelling strategy for a given biological system, otherwise inaccurate estimations and predictions may result, which could have serious consequences.

5. Acknowledgements

We would like to thank Parvathi Haridas for providing Figure 1. DCM would like to thank Oxford University Press for support through the Clarendon Fund, as well as Keble College, Oxford for support through the Sloane-Robinson award. This research is supported by the 2011 International Ex-

change Scheme funded by the Royal Society.

References

[1] Allred, D.C., 2010. Ductal carcinoma in situ: Terminology, classification and natural history. J. Nat. Cancer Inst. Monographs 41, 134–138.

[2] Ascolani, G., Badoual, M., Deroulers, C., 2013. Exclusion processes: short range correlations induced by adhesion and contact interactions. Phys Rev. E 87, 012792.

[3] Baker, R.E., Simpson, M.J., 2010. Correcting mean-field approximations for birth-death-movement processes. Phys. Rev. E. 82, 041905.

[4] Callaghan, T., Khain, E., Sander, L.M., Ziff, R.M., 2006. A stochastic model for wound healing. J. Stat. Phys. 122(5), 909–923.

[5] Cheng, G., Youssef, B.B., Markenscoff, P., Zygourakis, K., 2006. Cell population dynamics modulate the rates of tissue growth processes. Biophys. J. 90(3), 713–724.

[6] Codling, E.A., Plank, M.J., Benhamou, S., 2008. Random walk models in biology. J. R. Soc. Interface 5, 813–834.

[7] Dieckmann, U., Law, R., 2000. Relaxation Projections and the Method of Moments. Cambridge University Press. chapter 21. pp. 412–457.

[8] Dormann, S., Deutsch, A., 2002. Modeling of self-organized avascular tumor growth with a hybrid cellular automaton. In silico Biol. 2(3), 393–406.

[9] Ellner, S.P., Sasaki, A., Haraguchi, Y., Matsuda, H., 1998. Speed of invasion in lattice population models: pair-edge approximation. J. Math. Biol. 36, 469–484.

[10] Fisher, R.A., 1937. The wave of advance of advantageous genes. Ann. Eugenics 7, 353–369.

[11] Gatenby, R.A., Gawlinski, E.T., 1996. A reaction-diffusion model of cancer invasion. Cancer Res. 56, 5745.

[12] Khain, E., Lin, Y.T., Sander, L.M., 2011. Fluctuations and stability in front propagation. EPL 93, 28001.

[13] Kirkwood, J. G., 1935. Statistical mechanics of fluid mixtures. J. Chem. Phys. 3, 300-314.

[14] Kirkwood, J. G., Boggs, E. M., 1942. The radial distribution function in liquids. J. Chem. Phys. 10, 394-403.

[15] Kolmogorov, A., Petrovsky, I., Piscounov, N., 1937. Study of the diffusion equation with growth of the quantity of matter and its application to a biological problem. Bull. State Univ. Mos. , 1–25.

[16] Law, R., Murrell, D.J., Dieckmann, U., 2003. Population growth in space and time: spatial logistic equations. Ecology 84(1), 252–262.

[17] Lewis, M.A., 2000. Spread rate for a nonlinear stochastic invasion. J. Math. Biol. 41, 430–454.

[18] Liang, C.C., Park, A.Y., Guan, J.L., 2007. *In vitro* scratch assay: a convenient and inexpensive method for analysis of cell migration *in vitro*. Nat. Protoc. 2(2), 329–333.

[19] Mai, J., Kuzovkov, V.N., von Niessen, W., 1993. A theoretical stochastic model for the a+$1/2b_2 \rightarrow 0$ reaction. J. Chem. Phys. 98, 10017–10025.

[20] Mai, J., Kuzovkov, V.N., von Niessen W., 1994. A general stochastic model for the description of surface reaction systems. Physica A 203, 298–315.

[21] Maini, P.K., McElwain, D.L.S., Leavesley, D.I., 2004a. Traveling wave model to interpret a wound-healing cell migration assay for human peritoneal mesothelial cells. Tissue Eng. 10, 475–482.

[22] Maini, P.K., McElwain, D.L.S., Leavesley, D.I., 2004b. Travelling waves in a wound healing assay. Appl. Math. Model. 17, 575–580.

[23] Mani, S., Zygourakis, K., Markenscoff, P., 2002. Modeling wound healing: effects of cell heterogeneity and tissue structure, in: Proceedings of the Second Joint EMBS/BMES Conference, Houston, TX, USA.

[24] Markham, D.C., Simpson, M.J., Baker, R.E., 2013. Simplified method for including spatial correlations in mean-field approximations. Phys. Rev. E 87, 062702.

[25] Murray, J.D., 2002. Mathematical Biology I: An Introduction. Springer.

[26] Murrell, D.J., Dieckmann, U., Law, R., 2004. On moment closures for population dynamics in continuous space. J. Theor. Biol. 229, 421–432.

[27] Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 2007. Numerical Recipes: The Art of Scientific Computing. Cambridge University Press.

[28] Rodriguez, L.G., Wu, X., Guan, J.L., 2005. Methods in Molecular Biology, vol.294: Cell Migration: Developmental Methods and Protocols. Humana Press Inc., Totowa, NJ. chapter 3 Wound-healing assay. pp. 23–29.

[29] Sengers, B.G., Please, C.P., Oreffo, R.O.C., 2007. Experimental characterization and computational modelling of two-dimensional cell spreading for skeletal regeneration. J. R. Soc. Interface 4, 1107–1117.

[30] Sengers, B.G., Please, C.P., Taylor, M., Oreffo, R.O.C., 2009. Experimental-computational evaluation of human bone marrow stromal cell spreading on trabecular bone structures. Ann. Biomed. Eng. 37(6), 1165–1176.

[31] Sharkey, K.J., 2011. Deterministic epidemic models on contact networks: Correlations and unbiological terms. Theor. Pop. Biol. 79(4), 115–129.

[32] Simpson, M.J., Baker, R.E., 2011. Corrected mean-field models for spatially dependent advection-diffusion-reaction phenomena. Phys. Rev. E 83, 051922.

[33] Simpson, M.J., Binder, B.J., Haridas, P., Wood, B.K., Treloar, K.K., McElwain, D.L.S., Baker, R.E., 2013a. Experimental and modelling investigation of monolayer development with clustering. Bull. Math. Biol. 75, 871.

[34] Simpson, M.J., Merrifield, A., Landman, K.A., Hughes, B.D., 2007a. Simulating invasion with cellular automata: connecting cell-scale and population-scale properties. Phys Rev. E 76, 021918.

[35] Simpson, M.J., Treloar, K.K., Binder, B.J., Haridas, P., Manton, K.J., Leavesley, D.I., McElwain, D.L.S., Baker, R.E., 2013b. Quantifying the roles of cell motility and cell proliferation in a circular barrier assay. J. R. Soc. Interface 10, 20130007.

[36] Simpson, M.J., Zhang, D.C., Mariani, M., Landman, K.A., Newgreen, D.F., 2007b. Cell proliferation drives neural crest cell invasion of the intestine. Dev. Biol. 302, 553–568.

[37] Swanson, K.R., Bridge, C., Murray, J.D., Ellsworth, C.A.J., 2003. Virtual and real brain tumors: using mathematical modeling to quantify glioma growth and invasion. J. Neur. Sci. 216, 1–10.

[38] Todaro, G.J., Green, H., 1963. Quantitative studies of the growth of mouse embryo cells in culture and their development into established lines. J. Cell Bio. 17, 299–313.

[39] Tremel, A., Cai, A., Tirtaatmadja, N., Hughes, B.D., Stevens, G.W., Landman, K.A., O'Connor, A.J., 2009. Cell migration and proliferation during monolayer formation and wound healing. Chem. Eng. Sci. 64(2), 247–253.

[40] Valster, A., Tran, N.L., Nakada, M., Berens, M.E., Chan, A.Y., Symons, M., 2005. Cell migration and invasion assays. Methods 37, 208–215.

[41] Young, S.H., Rozengurt, N., Sinnett-Smith, J., Rozengurt, E., 2012. Rapid protein kinase D1 signaling promotes migration of intestinal epithelial cells. Am. J. Physiol. Gastrointest. Liver Physiol. 303, G356–G366.